

# Entropy measure for the quantification of upper quantile interdependence in multivariate distributions

Jhan Rodríguez<sup>a,\*</sup>, András Bárdossy<sup>a</sup>

<sup>a</sup>*Institute for Modelling Hydraulic and Environmental Systems, Universität Stuttgart*

---

## Abstract

We introduce a new measure of interdependence among the components of a random vector along the main diagonal of the vector copula, i.e. along the line  $u_1 = \dots = u_J$ , for  $(u_1, \dots, u_J) \in [0, 1]^J$ . Our measure is related to the Shannon entropy of a discrete random variable, hence we call it an “entropy index”. This entropy index is invariant with respect to marginal non-decreasing transformations and can be used to quantify the intensity of the vector components association in arbitrary dimensions. We show the applicability of our entropy index by an example with real data of 5 stock prices of the DAX index. In case the random vector possesses an extreme value copula, the index is shown to have as limit the extremal coefficient, which can be interpreted as the effective number of asymptotically independent components in the vector.

*Keywords:* Multivariate Interdependence, Entropy, Extremal Coefficient

---

## Introduction

The assessment of the intensity of tail dependence for multivariate data is an important task in several research areas, such as empirical finance, econometrics and atmospheric research. Additionally, the assessment of interdependence  
 5 among more than two random variables simultaneously has been indicated as relevant in research fields as diverse as weather forecasting, empirical finance and

---

\*Corresponding author

*Email address:* Jhan.Rodriguez@iws.uni-stuttgart.de (Jhan Rodríguez)

spike train analysis, in neuronal science Bárdossy and Pegram (2009); Staude et al. (2010); Dhaene et al. (2012).

It is useful to have some graphical tool to visualize the intensity of the association or dependence as one approaches the tail of the distribution. For example, the Chi-plot (Fisher and Switzer (2001)) has been used to pin down specific characteristics of tail behavior by Abberger (2005).

We present in this paper an index of association along the main diagonal of the copula of the distribution, i.e. along the line  $u_1 = \dots = u_J$ , for  $(u_1, \dots, u_J) \in [0, 1]^J$ . This index can be plotted to check for interdependence intensity at the uppermost quantiles of the distribution, as in Abberger (2005), but can also be readily applied to a random vector of dimension greater than two.

The rest of this paper is organized as follows: In section 1 we introduce the new association measure, to which we refer as an “entropy index”. In section 2 we apply the index to explore the type of dependence for 2, 3 and 4 dimensional marginal distributions of a real data set, and show how it can be used to evaluate goodness of tail fit for three different models fitted to the data. In section 3 we exhibit the relation of our entropy index with the extremal coefficient (see, for example, chapter 8 of Beirlant et al. (2004), and Schlather and Tawn (2003)), if the distribution of the analyzed vector is in the domain of attraction of an extreme value distribution. We end the paper with some conclusions and further interesting explorations of the entropy index.

## 1. The entropy index

We begin by recapitulating the "congregation measure" used by Bárdossy and Pegram (2009) and Bárdossy and Pegram (2012), for the sake of model validation. A modification of this measure constitutes the interdependence measure we introduce in this paper.

Let  $\mathbf{X} = (X_1, \dots, X_J)$  be a random vector with copula  $C$ , so that

$$F_{\mathbf{X}}(X_1, \dots, X_J) = C(F_1(X_1), \dots, F_J(X_J))$$

where  $F_{\mathbf{X}}$  is the probability distribution function of  $\mathbf{X}$  and  $F_1, \dots, F_J$  its marginal distribution functions. Our analysis focuses on the standardized random vectors

$$\mathbf{U} = (U_1, \dots, U_J) = (F_1(X_1), \dots, F_J(X_J))$$

Set a threshold percentile,  $b \in (0, 1)$ . Select a set of indexes  $(j_{i_1}, \dots, j_{i_K})$ , with  $1 \leq j_{i_1} < \dots < j_{i_K} \leq J$ . For the analysis of the components of  $\mathbf{U}$ , define binary random variables

$$\varsigma_b(j_{i_k}) = \begin{cases} 1, & U_{j_k} > b \\ 0, & U_{j_k} \leq b \end{cases} \quad (1)$$

This results in a discrete random vector,  $\varsigma_b = (\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K}))$ . The congregation measure introduced by Bárdossy and Pegram (2009) is defined to be the entropy of a sub-vector of  $\varsigma$ ,

$$H_b(U_{j_{i_1}}, \dots, U_{j_{i_K}}) = - \sum_{j_{i_1}, \dots, j_{i_K}} \Pr(\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K})) \log(\Pr(\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K}))) \quad (2)$$

That is, the measure is defined as the (Shannon) entropy of the joint distribution of the binary variables just defined. A higher value of this measure indicates less association, and vice versa.

Note that if the copula  $C$  is the independence copula,

$$C(u_1, \dots, u_J) = C^\sim(u_1, \dots, u_J) = \prod_{j=1}^{j=J} u_j$$

then the measure is constantly

$$H_b = -J(b \log(b) + (1-b) \log(1-b))$$

whereas if  $C$  is the co-monotonic copula,

$$C(u_1, \dots, u_J) = \min(u_1, \dots, u_J)$$

then the measure is also constant:

$$H_b = -(b \log(b) + (1-b) \log(1-b))$$

Between these two extremes lies the congregation measure, upon application to any given copula. Our entropy index is given by

$$S_b(\mathbf{U}) = \frac{H_b(U_{j_{i_1}}, \dots, U_{j_{i_K}})}{-(b \log(b) + (1-b) \log(1-b))} \quad (3)$$

This index quantifies the deviance from the totally dependent case. It is 1 in case of total dependence, and  $J$  in case of independence among the components of  $\mathbf{U}$ . Evidently, it can be used regardless of the dimension of  $\mathbf{U}$ , while keeping  
40 its interpretability as quantification of deviance from total independence.

In the following, we obviate the dependence on  $\mathbf{U}$  in order to make notation simpler.

An alternative, more general definition of the index, uses the so-called Tsallis entropy instead of the standard Shannon entropy. The Tsallis entropy includes an additional parameter  $\alpha \in (0, +\infty)$  and is defined by

$$H_b^\alpha = \frac{1}{\alpha - 1} \left( 1 - \sum_{j_{i_1}, \dots, j_{i_K}} \Pr(\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K}))^\alpha \right)$$

which reduces to  $H_b$  by letting  $\alpha \rightarrow 1$ . In this paper, we use the Tsallis entropy definition only as a technical tool in Appendix A, for proving a convergence  
45 result.

## 2. Example of applicability

We consider the stock prices of four components of the German DAX index, namely ADIDAS, ALLIANZ, BASF and BAYER, re-labeled in the following as components 1,2,3 and 4, respectively. The daily data spans the period  
50 from January 3th 2000 through June 30th 2014. The data used is available at [www.finanzen.de](http://www.finanzen.de).

For each stock  $j = 1, 2, 3, 4$ , we shall not consider the closing stock price at day  $t$ ,  $p_{t,j}$ , but rather the log-returns

$$r_{t,j} = 100 \times (\log(p_{t,j}) - \log(p_{t-1,j})) \quad (4)$$

This results in 4 time series, one for each stock. We use for our analysis of this 4-dimensional data set a model similar to that of Abberger (2005) and Dias and Embrechts (2004). The approach consists in fitting a time series model to each  
55 of the time series, independently, and then fitting a multivariate distribution to the resulting (presumably) iid vector formed out of the residuals of the time series.

To each of the four log-returns time series, we fit a GARCH(1,1) model. We obviate the  $j$  index in order to simplify notation, but the first equation below  
60 should read  $r_{t,j} = \mu_j + a_{t,j}$ , and so on. The model for each of the four time series is

$$r_t = \mu + a_t \quad (5)$$

$$a_t = \sigma_t \times \epsilon_t \quad (6)$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 a_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \quad (7)$$

Under the assumption that the  $\epsilon_t$  are iid with  $E(\epsilon_t) = 0$  and  $Var(\epsilon_t) = 1$ . Two typical assumptions for the so-called “standardized shocks”,  $\epsilon_t$ , are  $\epsilon_t \sim N(0, 1)$  and  $\sqrt{\frac{\nu}{\nu-2}} \times \epsilon_t \sim t_\nu$ , for  $\nu > 2$ . For details, the reader is referred,  
65 for example, to Tsay (2005). The idea of the GARCH model is to reproduce the clustering in variance, not explainable by linear time series models (like the ARMA model), often present in financial time series.

A GARCH(1,1) model was fitted independently to each log-returns time series using the garchFit function of package fGarch of the R statistical software  
70 (R Core Team (2014)). Estimation was performed using the Quasi maximum likelihood (QMLE) option, which is robust against miss-specification of the standardized shocks distribution.

The vectors of standardized shocks  $\epsilon_t = (\epsilon_{t,1}, \dots, \epsilon_{t,4})$ ,  $t = 1, \dots, 3686$  become now our object of study, or “observed data”. They are assumed to be  
75 (sufficiently) temporally independent, but there can be contemporaneous interdependence, with which we now deal. This contemporaneous interdependence, analogous to (Abberger (2005); Dias and Embrechts (2004)) can be modeled

by a copula model. Two models used in the literature are the Gaussian copula and the Student copula with unknown (i.e. to estimate) degrees of freedom,  $\nu$ .  
80 Their correlation matrices are also estimated in the process.

We fit in the following three models to the copula of random vector  $\epsilon$ . The first two are a Gaussian copula, and a Student copula. These copulas are fitted to the transformed standardized shocks,  $u_{t,j}$ ,  $j = 1, \dots, 4$ ,  $t = 1, \dots, 3686$ , obtained by

$$u_{t,j} = F_{n,j}(\epsilon_{t,j}) \quad (8)$$

where  $F_{n,j}$  stands for the empirical distribution function of the respective component,  $j$ , and is given by

$$F_{n,j}(\epsilon_{t,j}) = \frac{\#\{\epsilon_{s,j} : \epsilon_{s,j} \leq \epsilon_{t,j}\}}{3687}$$

The third model explored here is not a copula model. We represent the density of  $\epsilon$  by a mixture of five multivariate normal distributions

$$f(\epsilon) = \sum_{k=1}^5 w_k g_k(\epsilon)$$

where each Normal distribution is allowed to have its own mean vector and covariance matrix; such mean vectors, covariance matrices and the weights ( $w_1, \dots, w_k$ ) are estimated on the basis of the available data.

We used the R software to fit all the respective model parameters: Function  
85 `fitCopula` of package *copula* to fit the copula models parameters; and function `init.EM` of package *EMcluster* to fit the mixture model, which estimates the mixture parameters by the Expectation Maximization algorithm. The fitted degrees of freedom for the Student copula model were 7.50, which are those of a model with non-negligible tail dependence.

90 We proceed now to the analysis of the joint association of 2,3 and 4 dimensional joint marginals of vector  $\epsilon$ , both of the observed data and of data simulated from the three fitted models. To this end, we use the first 2, 3 and 4 components of  $\epsilon$ , respectively. Our analysis consists in computing the entropy index defined in section 1 for increasing quantile thresholds,  $b \in (0.85, 1.00)$ ,

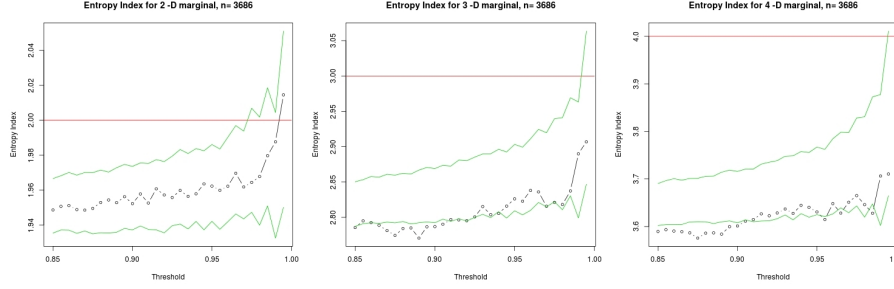


Figure 1: Entropy index computed for selected quantile thresholds: Data (Black dotted line) and Monte Carlo confidence interval for data coming from the fitted Gaussian copula.

at the upper part of the distribution. Specifically, we use threshold values  $b = .850, .855, \dots, .995$ .

We shall see how the degree of association of each of the models along the line  $u_1 = \dots = u_4 = b$  is, as compared to that of the observed data. This gives us an idea of the adequacy the modeled interdependence, for the 2,3 and 4-dimensional marginals, as one focuses on the uppermost part of the distribution.

In figure 1 we show the entropy index computed for the observed data at the indicated thresholds,  $b$ , given by the black lines and points. The green lines added correspond to a 95% confidence interval for data obtained from the fitted Gaussian copula model. The confidence interval is based on the generation of 500 data sets, each of size 3686, of the fitted Gaussian copula model, and the computation of the entropy index for the thresholds  $b = .850, .855, \dots, .995$ , as had been done for observed data.

From figure 1, we see that the type of association in the observed data, as represented by the entropy index, is similar in the two-dimensional marginals to that of the Gaussian copula. Data association is however systematically stronger, for the 3 and 4 dimensional distributions considered: the black line sticks to the bottom of the confidence interval, sometimes even stepping out of it.

This is another warning about the need to validate multivariate statistical models by statistics that consider more than two components at a time, when

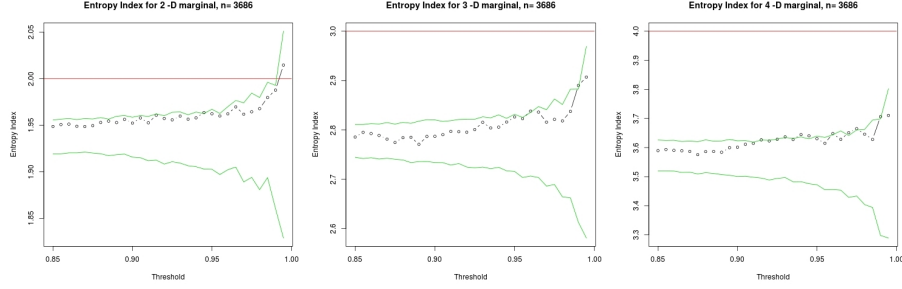


Figure 2: Entropy index computed for selected quantile thresholds: Data (Black dotted line) and Monte Carlo confidence interval for data coming from the fitted Student-t copula.

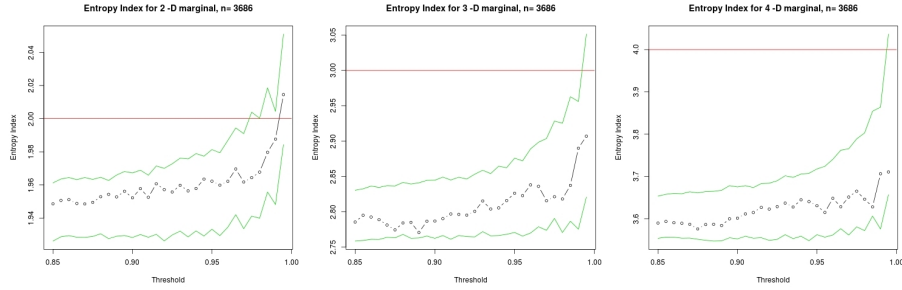


Figure 3: Entropy index computed for selected quantile thresholds: Data (Black dotted line) and Monte Carlo confidence interval for data coming from the fitted mixture of Normal distributions.

joint interaction among more than two components is relevant for the problem at hand (cf. Bárdossy and Pegram (2009, 2012); Rodríguez and Bárdossy (2014)).

In figure 2, we show the same type of plot as before, but for the student copula model. We note that the interdependence is somewhat exaggerated, even  
120 for the 2-dimensional marginal considered. This exaggeration is clearer for the 4-dimensional joint distribution. Even if the asymptotic tail dependence is not adequate for high (though not extreme) quantiles of the joint distributions shown.

125 The same procedure as above is repeated with the mixture model. The result is shown in figure 3. Note that the black line, corresponding to the measure of



interdependence for data, lies roughly at the center of the confidence interval, for each of the joint 2, 3 and 4-dimensional marginals presented.

This is an indication that, in terms of the type of association measured by  
 130 our entropy index, the mixture model is a more realistic representation of the process, as compared with the other two models. That it is a better representation of the variables interdependence at the section of the distribution just before the extreme value region, as one *approaches* that region.

This better fit is not much of a surprise, since 5 mixture components provide  
 135 considerable modeling flexibility. Our point here is not to favor the a specific, over-parameterized model, but to show how one can notice important deficiencies in the fit of a given model in a  $d \geq 2$  -dimensional setting.

Sometimes finance data, like the one here shown, is subject to tail dependence (cf. Abberger (2005)). In that case, an even better model would be a  
 140 mixture of Student distributions, with degrees of freedom higher than 7, each. In this way, the joint association for high quantiles would not be exaggerated, while the asymptotic tail dependence will not be zero, as in the case of a mixture of multivariate normal distributions.

### 3. Relation with extremal coefficient

An extreme value copula is the copula of an extreme value distribution,  $G$ , and can be characterized by the following stability condition: A copula on  $[0, 1]^J$  is of the extreme value type if, and only if,

$$C_G^s(u_1, \dots, u_J) = C_G(u_1^s, \dots, u_J^s) \quad (9)$$

145 for all  $s > 0$ .

For  $b \in [0, 1]$ , extreme value copulas fulfill the relation

$$C_G(b, \dots, b) = b^\theta \quad (10)$$

for some  $\theta \in [1, J]$ . This parameter  $\theta$  receives the name of *extremal coefficient* (see, for example chapter 8 of Beirlant et al. (2004)). It can be thought of

as the asymptotic “effective number of independent variables” (Schlather and Tawn (2003)) of  $\mathbf{X}$ . Its value lies between 1 and  $J$ , in case of asymptotic perfect  
150 dependence and complete independence, respectively.

Assume that the distribution of  $\mathbf{X} = (X_1, \dots, X_J)$  is in the domain of attraction of an extreme value distribution,  $G$ , i.e.  $F_{\mathbf{X}} \in D(G)$ . Since (Beirlant et al. (2004), p. 282)

$$b \rightarrow 1^- \Rightarrow C_F(b, \dots, b) \rightarrow C_G(b, \dots, b) \quad (11)$$

one then has that:

$$C_{F_{\mathbf{X}}}(b, \dots, b) \rightarrow b^\theta \quad (12)$$

so that, also for the copula of  $\mathbf{X}$ , interdependence along the main diagonal of the copula is determined by the extremal coefficient. The extremal coefficient is the same in both cases.

The limit of our entropy index is precisely the extremal coefficient of the copula of  $\mathbf{X}$ . Namely,

$$\lim_{b \rightarrow 1^-} S_b(\mathbf{U}) = \theta \quad (13)$$

for  $\mathbf{U} = (U_1, \dots, U_J)$ ,  $U_j = F_j(X_j)$ , and  $F_j$  the marginal distribution of  $X_j$ , for  
155  $j = 1, \dots, J$ .

The proof of (13) is rather technical, so it is relegated to the appendix.

The following is thus an application of the entropy index: For a random vector  $\mathbf{X}$  whose distribution is in the domain of attraction of an extreme value distribution, we can explore how “fast” its asymptotic effective number of inde-  
160 pendent components is approached.

As an example, see figure 4, where the entropy index is computed for two different copulas, for thresholds  $b = .8, .9, .95, .99, .995$ . Figure 4 is the result of computing the entropy index for the mentioned thresholds to a simulated sample of size  $n = 10^6$ , for each distribution; so figure 4 is a good approximation to a  
165 figure based on exact, analytic computations.

The green line corresponds to a Gumbel 3-dimensional copula,

$$C(u_1, u_2, u_3) = \exp \left\{ - \left[ (-\log(u_1))^\xi + \dots + (-\log(u_3))^\xi \right]^{\frac{1}{\xi}} \right\} \quad (14)$$

for which the extremal coefficient is  $\theta = 3^{\frac{1}{\xi}}$ .

The black line at figure 4 corresponds to a student copula with  $\nu$  degrees of freedom and correlation matrix  $\rho \in \mathbb{R}^{3 \times 3}$ , for which the extremal coefficient is given by

$$\theta = \sum_{j=1}^3 T_{2, \nu+1, \rho_{-j, -j}} \left( \frac{\sqrt{\nu+1}}{\sqrt{1-\rho_{i,j}}} (1-\rho_{i,j}), i \neq j \right) \quad (15)$$

where  $T_{2, \nu+1, \rho_{-j, -j}}$  stands for the 2-dimensional Student cumulative distribution function with  $\nu+1$  degrees of freedom and dispersion matrix  $\rho_{-j, -j}$ . In turn,  $\rho_{-j, -j}$  is the  $2 \times 2$  matrix resulting from removing row and column  $j$  from  $\rho$ .

170 That (15) contains the extremal coefficient in question can be readily seen from Theorem 2.3, equation 2.8, of Nikoloulopoulos et al. (2009).

Parameters  $\xi$ ,  $\nu$  and  $\rho$  were selected in such a way that, for both distributions, the extremal coefficient is  $\theta = 2$ . The specific parameters used can be found in the appendix.

175 In spite of having the same asymptotic “effective number of independent components”, the association among the components of the Student copula is systematically stronger (in terms of the entropy index) before reaching that limit.

This is an additional support in favor of our entropy index as a means of  
180 analyzing interdependence carefully at the uppermost part of the distribution. This detailed analysis can be useful for goodness of fit purposes.

#### 4. Conclusion and future work

We have introduced a tool that is useful for exploring the intensity of association at the joint upper quantiles of a multivariate distribution. This tool  
185 is not limited to 2-dimensional distributions. We can have a measure of how strong or weak is the association just before the extreme value case, where the intensity of dependence may also be important for some applications.

The limit of the entropy index, in case of a vector having a distribution in the domain of attraction of an extreme value distribution, is the extremal  
190 coefficient,  $\theta$ . This coefficient can be interpreted as the asymptotic effective

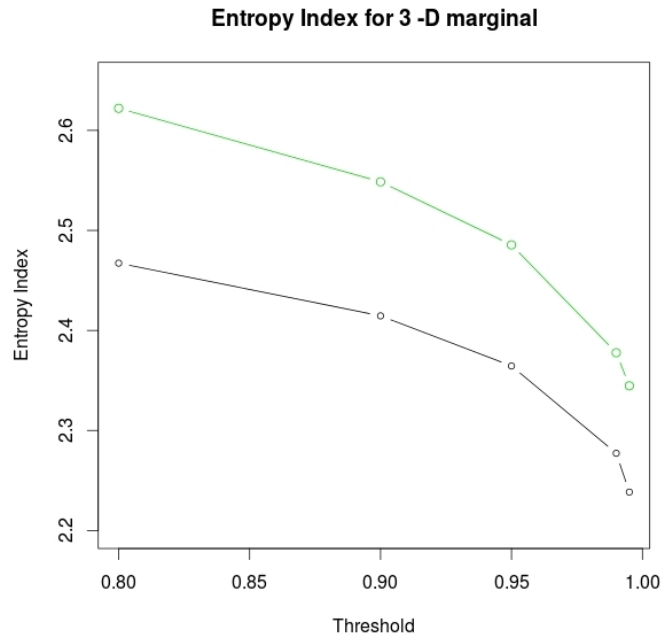


Figure 4: Comparison of entropy index for the 3-dimensional Gumbel copula (green) and Student copula (black). Thresholds are  $b = .8, .9, .95, .99, .995$ . Both distributions have extremal coefficient  $\theta = 2$ , towards which they approach. However, note that the Student copula approaches its asymptotic number of effectively independent components faster.

number of independent components in the random vector. So, using the entropy index presented in this article, we can have an idea of how fast this asymptotic value is reached by the vector components.

### *Acknowledgments*

195 This paper stems from the Ph.D research of the first author, which was funded by a scholarship of the German Academic Exchange Service (DAAD). This Ph.D work was carried out within the framework of the ENWAT program at the University of Stuttgart.

### **References**

- 200 Abberger, K., 2005. A simple graphical method to explore tail-dependence in stock-return pairs. *Applied Financial Economics* 15, 43–51.
- Bárdossy, A., Pegram, G., 2009. Copula based multisite model for daily precipitation simulation. *Hydrology and Earth System Sciences Discussions* 6, 4485–4534. URL: [http://www.hydrol-earth-syst-sci-discuss.net/6/4485/](http://www.hydrol-earth-syst-sci-discuss.net/6/4485/2009/)  
205 2009/, doi:10.5194/hessd-6-4485-2009.
- Bárdossy, A., Pegram, G., 2012. Multiscale spatial recorrelation of RCM precipitation to produce unbiased climate change scenarios over large areas and small. *Water Resources Research* 48. URL: <http://www.agu.org/pubs/crossref/2012/2011WR011524.shtml>, doi:10.1029/2011WR011524.
- 210 Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J., 2004. *Statistics of extremes: theory and applications*. John Wiley & Sons.
- Dhaene, J., Linders, D., Schoutens, W., Vyncke, D., 2012. The herd behavior index: A new measure for the implied degree of co-movement in stock markets. *Insurance: Mathematics and Eco-*  
215 *nomics* 50, 357 – 370. URL: <http://www.sciencedirect.com/science/article/pii/S0167668712000066>, doi:<http://dx.doi.org/10.1016/j.insmatheco.2012.01.005>.

- Dias, A., Embrechts, P., 2004. Dynamic copula models for multivariate high-frequency data in finance. Manuscript, ETH Zurich .
- 220 Fisher, N., Switzer, P., 2001. Graphical assessment of dependence: Is a picture worth 100 tests? *The American Statistician* 55, 233–239.
- Nikoloulopoulos, A.K., Joe, H., Li, H., 2009. Extreme value properties of multivariate t copulas. *Extremes* 12, 129–148.
- R Core Team, 2014. R: A Language and Environment for Statistical Computing. Vienna, Austria. URL: <http://www.R-project.org/>.
- 225 Rodríguez, J., Bárdossy, A., 2014. Beyond correlation in spatial statistics modeling (submitted). ArXiv e-prints [arXiv:1406.2501](https://arxiv.org/abs/1406.2501).
- Schlather, M., Tawn, J.A., 2003. A dependence measure for multivariate and spatial extreme values: Properties and inference. *Biometrika* 90, 139–156.
- 230 Staude, B., Rotter, S., Grün, S., 2010. Cubic: cumulant based inference of higher-order correlations in massively parallel spike trains. *Journal of Computational Neuroscience* 29, 327–350. URL: <http://dx.doi.org/10.1007/s10827-009-0195-x>, doi:10.1007/s10827-009-0195-x.
- Tsay, R.S., 2005. Analysis of financial time series. volume 543. John Wiley & Sons.
- 235

## Appendix A. Proof of convergence to Extremal coefficient

We present here the proof of equation (13).

We assume in the following that random vector  $\mathbf{U} \in [0, 1]^J$  has an extreme value copula, so that  $C(b, \dots, b) = b^\theta$ . By virtue of equation (11), the argument

240 could be repeated for  $\mathbf{U} \in [b_0, 1]^J$ , with  $b_0$  sufficiently close to 1, discarding the hypothesis of an extreme value copula, but assuming the original random vector in the domain of attraction of an extreme value distribution.

Let random vector  $\mathbf{U} \in [0, 1]^J$  have a copula,  $C$ , of the extreme value type. We show in this section that, for any  $\alpha > 1$ ,

$$\lim_{b \rightarrow 1^-} T_b^\alpha(\mathbf{U}) = \theta \quad (\text{A.1})$$

Then, by letting  $\alpha \rightarrow 1^+$ , we shall have that  $\lim_{b \rightarrow 1^-} S_b(\mathbf{U}) = \theta$ .

To simplify notation, dependence of  $S_b$  on random vector  $\mathbf{U}$  is taken for granted in the following, so that, for example,  $S_b := S_b(\mathbf{U})$ , and so on.

To prove equation (A.1) we shall provide “sandwich” functions  $g_1(b)$  and  $g_2(b)$ , such that  $g_1(b) \leq T_b^\alpha \leq g_2(b)$  for all  $b$ ,  $b_0 < b < 1$  (for some sufficiently large  $b_0$ ). Of these auxiliary functions, it will be easy to show that

$$\lim_{b \rightarrow 1^-} g_1(b) = \lim_{b \rightarrow 1^-} g_2(b) = \theta$$

whence, necessarily, one must have (A.1).

Of all the probabilities appearing at (2), the most straightforward to identify is

$$\Pr(\varsigma_b(1) = 0, \dots, \varsigma_b(J) = 0) = C(b, \dots, b)$$

if we have the function defining  $C$ . Under the assumption that  $\mathbf{U}$  has an extreme value copula, this probability is simply  $C(b, \dots, b) = b^\theta$ . The other probabilities can be very difficult to evaluate in terms of the original copula,  $C$ . So we shall try to use this value to our convenience.

Note that if  $b \rightarrow 1^-$ , then

$$\frac{(1 - b^\alpha)}{(1 - b)^\alpha} \rightarrow +\infty$$

and hence as we approach 1 from below, it is right to assume that above certain  $0 < b_0 < 1$ , one has  $(1 - b^\alpha) - (1 - b)^\alpha > 0$ .

For any  $b \in [0, 1]$ , one has  $\Pr(\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K})) \leq (1 - b)$ , regardless of whether  $\varsigma(j_{i_k}) = 0$  or  $\varsigma(j_{i_k}) = 1$  for each  $j_{i_k}$  in the index set. Then

$$1 - (\Pr(\varsigma_b(j_{i_1}) = 0, \dots, \varsigma_b(j_{i_K}) = 0)^\alpha + (2^J - 1)(1 - b)^\alpha) \leq 1 - \sum_{j_{i_1}, \dots, j_{i_K}} \Pr(\varsigma_b(j_{i_1}), \dots, \varsigma_b(j_{i_K}))^\alpha$$

where the term  $(2^J - 1)(1 - b)^\alpha$  accounts for the remaining  $(2^J - 1)$  probability values, apart from  $\Pr(\varsigma_b(j_{i_1}) = 0, \dots, \varsigma_b(j_{i_K}) = 0)$ .

Thus, using the extreme value copula assumption from equation (12), one can define for  $b > b_0$  the function

$$g_1(b) := \frac{1 - (\Pr(\varsigma_b(j_{i_1}) = 0, \dots, \varsigma_b(j_{i_K}) = 0)^\alpha + (2^J - 1)(1 - b)^\alpha)}{(1 - b^\alpha) - (1 - b)^\alpha} = \frac{1 - (b^{\theta\alpha} + (2^J - 1)(1 - b)^\alpha)}{(1 - b^\alpha) - (1 - b)^\alpha} \leq T_b^\alpha \quad (\text{A.2})$$

255 Note that we require  $b > b_0$  to ensure that  $(1 - b^\alpha) - (1 - b)^\alpha > 0$ .

On the other hand, for  $b$  sufficiently large,

$$g_2(b) := \frac{1 - b^{\theta\alpha}}{(1 - b^\alpha) - (1 - b)^\alpha} = \frac{1 - \Pr(\varsigma_b(j_{i_1}) = 0, \dots, \varsigma_b(j_{i_K}) = 0)^\alpha}{(1 - b^\alpha) - (1 - b)^\alpha} \geq T_b^\alpha \quad (\text{A.3})$$

Then, for any  $\alpha > 1$ , and  $b$  such that  $b > b_0$ , one has

$$g_1(b) \leq T_b^\alpha \leq g_2(b)$$

Concerning limits, applying L'Hôpital's rule,

$$\lim_{b \rightarrow 1} g_1(b) = \lim_{b \rightarrow 1} \frac{1 - (b^{\theta\alpha} + (2^J - 1)(1 - b)^\alpha)}{(1 - b^\alpha) - (1 - b)^\alpha} = \lim_{b \rightarrow 1} \frac{\alpha(1 - b)^{\alpha-1}(2^J - 1) - \alpha\theta b^{\alpha\theta-1}}{\alpha(1 - b)^{\alpha-1} - \alpha b^{\alpha-1}} = \theta$$

and hence  $\lim_{b \rightarrow 1^-} g_1(b) = \theta$ .

Similarly, one can see using L'Hôpital's rule that

$$\lim_{b \rightarrow 1^-} g_2(b) = \theta$$

Hence, one must have, for any  $\alpha > 1$ , that

$$\lim_{b \rightarrow 1^-} T_b^\alpha = \theta \quad (\text{A.4})$$

as we wanted to show.

Now, the Tsallis entropy fulfills  $H_b^\alpha \rightarrow H_b$ , as  $\alpha \rightarrow 1^+$ . Hence one has

$$\theta = \lim_{\alpha \rightarrow 1^+} \left( \lim_{b \rightarrow 1^-} T_b^\alpha \right) = \lim_{b \rightarrow 1^-} S_b \quad (\text{A.5})$$

as we wanted to show in this part of the appendix.



## Appendix B. Parameters used for example of section 3

260 We present here the parameters used for the example at section 3. These parameters were selected in such a way that for the extremal coefficient,  $\theta$ , one has  $\theta = 2$ .

For the Gumbel copula, the dependence parameter used was  $\xi = \frac{\log(3)}{\log(2)} \approx 1.585$ .

For the Student copula,  $\nu = 2.76733$ , and

$$\rho = \begin{pmatrix} 1 & .767 & .759 \\ .767 & 1 & .624 \\ .759 & .624 & 1 \end{pmatrix}$$